



Europäisches
Patentamt

European
Patent Office

Office européen
des brevets

RECD 22 APR 2004

SEARCHED

INDEXED

Bescheinigung

Certificate

Attestation

Die angehefteten Unterlagen stimmen mit der ursprünglich eingereichten Fassung der auf dem nächsten Blatt bezeichneten europäischen Patentanmeldung überein.

The attached documents are exact copies of the European patent application described on the following page, as originally filed.

Les documents fixés à cette attestation sont conformes à la version initialement déposée de la demande de brevet européen spécifiée à la page suivante.

Patentanmeldung Nr. Patent application No. Demande de brevet n°

03007274.8

**PRIORITY
DOCUMENT**
SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH RULE 17.1(a) OR (b)

BEST AVAILABLE COPY

Der Präsident des Europäischen Patentamts;
Im Auftrag

For the President of the European Patent Office

Le Président de l'Office européen des brevets
p.o.

R C van Dijk



Anmeldung Nr:
Application no.: 03007274.8
Demande no:

Anmeldetag:
Date of filing: 31.03.03
Date de dépôt:

Anmelder/Applicant(s)/Demandeur(s):

DeveloGen Aktiengesellschaft für
entwicklungsbiologische Forschung
Rudolf-Wissell-Strasse 28
37079 Göttingen
ALLEMAGNE

Bezeichnung der Erfindung/Title of the invention/Titre de l'invention:
(Falls die Bezeichnung der Erfindung nicht angegeben ist, siehe Beschreibung.
If no title is shown please refer to the description.
Si aucun titre n'est indiqué se referer à la description.)

Use of a DG931 protein product for preventing and treating obesity and/or
diabetes and/or metabolic syndrome

In Anspruch genommene Priorität(en) / Priority(ies) claimed /Priorité(s)
revendiquée(s)
Staat/Tag/Aktenzeichen/State/Date/File no./Pays/Date/Numéro de dépôt:

Internationale Patentklassifikation/International Patent Classification/
Classification internationale des brevets:

A61K38/00

Am Anmeldetag benannte Vertragstaaten/Contracting states designated at date of
filing/Etats contractants désignées lors du dépôt:

AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HU IE IT LU MC NL
PT SE SI SK TR LI

31. März 2003

- 1 -

**USE OF A DG931 PROTEIN PRODUCT FOR PREVENTING AND
TREATING OBESITY AND/OR DIABETES AND/OR METABOLIC
SYNDROME**

5

Description

This invention relates to the use of DG931 proteins, to the use of polynucleotides encoding these, and to the use of effectors / modulator thereof in the diagnosis, study, prevention, and treatment of obesity and / or diabetes mellitus and/or metabolic syndrome.

Many human proteins serve as pharmaceutically active compounds. Several classes of human proteins that serve as such active compounds include hormones, cytokines, cell growth factors, and cell differentiation factors. Most proteins that can be used as a pharmaceutically active compound fall within the family of secreted proteins. Secreted proteins are generally produced within cells at rough endoplasmic reticulum, are then exported to the golgi complex, and then move to secretory vesicles or granules, where they are secreted to the exterior of the cell via exocytosis. Examples for commercially used secreted proteins are human insulin, thrombolytic agents, interferons, interleukins, erythropoietins, colony stimulating factors, human growth hormone, transforming growth factor beta, tissue plasminogen activator, erythropoietin, and various other proteins. Receptors of secreted proteins, which are membrane-bound proteins, also have potential as therapeutic or diagnostic agents.

It is, therefore, important for developing new pharmaceutical compounds to identify secreted proteins that can be tested for activity in a variety of animal models. Thus, in light of the pervasive role of secreted proteins in human physiology, a need exists for identifying and characterizing novel functions for human secreted proteins and the genes that encode them. This knowledge will allow one to detect, to treat, and to prevent medical diseases, disorders, and/or conditions by using secreted proteins or the genes that encode them.

Obesity is one of the most prevalent metabolic disorders in the world. It is still a poorly understood human disease that becomes as a major health problem more and more relevant for western society. Obesity is defined as a body weight more than 20% in excess of the ideal body weight, frequently resulting in a significant impairment of health. Obesity may be measured by body mass index, an indicator of adiposity or fatness. Further parameters for defining obesity are waist circumferences, skinfold thickness and bioimpedance (see, *inter alia*, Kopelman (1999), *loc. cit.*). It is associated with an increased risk for cardiovascular disease, hypertension, diabetes mellitus type II, hyperlipidaemia and an increased mortality rate.

Obesity is influenced by genetic, metabolic, biochemical, psychological, and behavioral factors and can be caused by different reasons such as non-insulin dependent diabetes, increase in triglycerides, increase in carbohydrate bound energy and low energy expenditure. As such, it is a complex disorder that must be addressed on several fronts to achieve lasting positive clinical outcome. Since obesity is not to be considered as a single disorder but as a heterogeneous group of conditions with (potential) multiple causes, it is also characterized by elevated fasting plasma insulin and an exaggerated insulin response to oral glucose intake (Koltermann J., (1980) *Clin. Invest.* 65, 1272-1284). A clear involvement of obesity in type 2 diabetes mellitus can be confirmed (Kopelman P.G., (2000) *Nature* 404, 635-643).

Insulin amongst other hormones plays a key role in the regulation of the fuel metabolism. High blood glucose levels stimulate the secretion of insulin by pancreatic beta-cells. Insulin leads to the storage of glycogen and triglycerides and to the synthesis of proteins. The entry of glucose into muscles and adipose cells is stimulated by insulin. In patients who suffer from diabetes mellitus either the amount of insulin produced by the pancreatic islet cells is to low (diabetes type I or insulin dependent diabetes mellitus - IDDM) or liver and muscle cells loose their ability to respond to normal blood insulin levels (insulin resistance). In the next stage pancreatic cells become unable to produce sufficient amounts of insulin (diabetes type II or non insulin dependent diabetes mellitus - NIDDM).

The concept of 'metabolic syndrome' (syndrome x, insulin-resistance syndrome, deadly quartet) was first described 1966 by Camus and reintroduced 1988 by Reaven (Camus JP, 1966, Rev Rhum Mal Osteoartic 33(1):10-14; Reaven et al. 1988, 5 Diabetes, 37(12):1595-1607). Today metabolic syndrome is commonly defined as clustering of cardiovascular risk factors like hypertension, abdominal obesity, high blood levels of triglycerides and fasting glucose as well as low blood levels of HDL cholesterol. Insulin resistance greatly increases the risk of developing the metabolic syndrome (Reaven, 2002, Circulation 106(3): 286-288). The metabolic syndrome 10 often precedes the development of type II diabetes and cardiovascular disease (McCook, 2002, JAMA 288:2709-2716). The control of blood lipid levels and blood glucose levels is the essential for the treatment of the metabolic syndrome (see, for example, Santomauro A. T. et al., (1999) Diabetes, 48(9):1836-1841).

15 The molecular factors regulating food intake and body weight balance are incompletely understood. Even if several candidate genes have been described which are supposed to influence the homeostatic system(s) that regulate body mass/weight, like leptin or the peroxisome proliferator-activated receptor-gamma co-activator, the distinct molecular mechanisms and/or molecules influencing obesity or body weight/body mass regulations 20 are not known.

The cysteine rich secreted protein (CRISP) family is composed of over 70 members that are found in many phyla of organisms, including: vertebrates, plants, fungi, yeast, and insects. The cysteine rich secreted protein of this invention, referred to as DG931, is a 25 member of this family. No functional data are available in the scientific prior art for this protein. However, the protein and possible functions are disclosed in several patent applications. For example, DG931 is described as ,novel protease-inhibitor like protein', useful in disease diagnosis and treatment, particularly of diseases associated with signal transduction (see WO 01/21651 A2). Patent application WO 01/62928 30 A2 describes the same protein as ,KIAA1233-like' or ,STE20-like' or ,trypsin-inhibitor like protein' for treating a pathology related to aberrant expression or aberrant physiological condition, for example, haemetopoietic, cancer,

immunological, tumour, neurodegenerative, AIDS, or fertility disorders. Applications WO 02/08288 A2 and WO 02/22802 A1 describe DG931 as being useful in disease diagnosis and treatment, particularly of treating and diagnosing human tumors. To the best of our knowledge, no disclosure has been made in the 5 prior art that describes a function of DG931 in metabolic disorders such as obesity, diabetes, or metabolic syndrome.

10 Secreted proteins are a major target for drug action and development. Accordingly, it is valuable to the field of pharmaceutical development to identify and characterize novel functions for secreted proteins. The present invention advances the state of the art by providing previously unknown functions for a human secreted protein that has homology to cystein-rich secreted proteins.

15 Accordingly, the present invention relates to a secreted protein with novel functions in the human metabolism. The present invention discloses specific genes and proteins encoded thereby involved in the regulation of metabolism, especially in the disorders of the metabolic syndrome, obesity, diabetes mellitus, coronary heart disease, eating disorder, cachexia, hypertension, hypercholesterolemia (dyslipidemia), and/or gallstones.

20 In this invention, we used a screen for secreted factors expressed in developing mammalian (mouse) pancreas, as described in more detail in the Examples section (see Example 1). This screen identified DG931 as secreted factors expressed in developing mouse pancreas. The present invention describes mammalian DG931 proteins and the 25 polynucleotides encoding these, in particular human DG931, as being involved in the conditions mentioned above.

30 The present invention relates to DG931 polynucleotides encoding polypeptides with novel functions in mammalian metabolism, including body-weight regulation, energy homeostasis, obesity, and diabetes, fragments of said polynucleotides, polypeptides encoded by said polynucleotides or fragments thereof. The invention also relates to vectors, host cells, antibodies, and recombinant methods for producing the polypeptides and polynucleotides of the invention. The invention also relates to effectors / modulators

of DG931 polynucleotides and / or polypeptides, e.g. antibodies, biologically active nucleic acids, such as antisense molecules, RNAi molecules or ribozymes, aptamers, peptides or low-molecular weight organic compounds recognizing said polynucleotides or polypeptides.

5 DG931 homologous proteins and nucleic acid molecules coding therefore are obtainable from vertebrate species. Particularly preferred are nucleic acids encoding the human DG931 protein and variants thereof.

The invention particularly relates to a nucleic acid molecule encoding a polypeptide
10 contributing to regulating the energy homeostasis and the mammalian metabolism,
wherein said nucleic acid molecule comprises

- (a) the nucleotide sequence of human DG931 (SEQ ID NO: 1) and/or a sequence complementary thereto,
- (b) a nucleotide sequence which hybridizes at 50°C in a solution containing 1 x SSC
15 and 0.1% SDS to a sequence of (a),
- (c) a sequence corresponding to the sequences of (a) or (b) within the degeneration of the genetic code,
- (d) a sequence which encodes a polypeptide which is at least 85%, preferably at least 90%, more preferably at least 95%, more preferably at least 98% and
20 up to 99,6% identical to the amino acid sequences of the human DG931 protein (SEQ ID NO: 2),
- (e) a sequence which differs from the nucleic acid molecule of (a) to (d) by mutation and wherein said mutation causes an alteration, deletion, duplication and/or premature stop in the encoded polypeptide or
- 25 (f) a partial sequence of any of the nucleotide sequences of (a) to (e) having a length of 15-25 bases, preferably 25-35 bases, more preferably 35-50 bases and most preferably at least 50 bases.

In one embodiment of the invention, the function of the mammalian DG931 in
30 metabolism was further validated by analyzing the expression of the transcripts in different tissues and by analyzing the role in adipocyte differentiation.

Expression profiling studies (see Examples for more detail) confirm the particular relevance of DG931 as regulator of energy metabolism in mammals. Taqman analysis revealed that DG931 is expressed in several mammalian tissues, with highest expression levels in heart, lung, kidney, testis and muscle. In addition,

5 DG931 is highly expressed in metabolic active tissue such as white adipose tissue (WAT) and brown adipose tissue (BAT) compared to other tissue types in wild type mouse as depicted in FIGURE 2A. BAT is a well characterized tissue which is well developed in newborn mammals, including humans. One important task of BAT is to generate heat and maintain body temperature homeostasis in newborn. Thus, an
10 expression of DG931 protein in adipose tissues is confirming a role in the regulation of metabolism, particularly energy homeostasis and thermogenesis.

In one embodiment of this invention, we used mouse models of insulin resistance and/or diabetes, such as mice carrying gene knockouts in the leptin pathway (for

15 example, *ob/ob* (leptin) or *db* (leptin receptor/ligand) mice) to study the expression of DG931. Such mice develop typical symptoms of diabetes, show hepatic lipid accumulation and frequently have increased plasma lipid levels (see Bruning et al, 1998, Mol. Cell. 2:449-569). We found, for example, that the expression of DG931 is strongly upregulated in metabolic active tissue (BAT, WAT) and in liver tissue in
20 genetically induced obese mice (*ob/ob*) compared to fasted mice (see FIGURE 2B). These data further support an essential role of DG931 in the regulation of the mammalian metabolism, particularly in processes related to diabetes and obesity.

In one embodiment of the invention, expression of DG931 mRNA was also

25 examined in susceptible wild type mice (for example, C57Bl/6) that show symptoms of diabetes, lipid accumulation, and high plasma lipid levels, if fed a high fat diet. In those mice, the expression of DG931 is significantly upregulated in several tissues, including muscle, BAT, and WAT supporting that DG931 is involved in the regulation of mammalian metabolism (see FIGURE 2C).

30

In another embodiment of the invention, we show (see Examples and FIGURES 2 and 3) that the DG931 protein has to be significantly downregulated in order for the

preadipocytes to differentiate into mature adipocyte. With regard to changes in expression intensity during the differentiation of preadipocytes to adipocytes, a strong reduction in relative signal intensity can be observed for DG931 during the *in vitro* differentiation program of 3T3-L1 (see Figure 2D). The DG931 protein in 5 preadipocytes has the potential to inhibit adipose differentiation at a very early stage (FIGURE 2D). Therefore, the DG931 protein might play an essential role in adipogenesis. The results are suggesting a role as modulator (for example, inhibitor) of adipogenesis.

10 Microarrays are analytical tools routinely used in bioanalysis. A microarray has molecules distributed over, and stably associated with, the surface of a solid support. The term "microarray" refers to an arrangement of a plurality of polynucleotides, polypeptides, antibodies, or other chemical compounds on a substrate. Microarrays of polypeptides, polynucleotides, and/or antibodies have been developed and find use in a 15 variety of applications, such as monitoring gene expression, drug discovery, gene sequencing, gene mapping, bacterial identification, and combinatorial chemistry. One area in particular in which microarrays find use is in gene expression analysis (see Example 4). Array technology can be used to explore the expression of a single polymorphic gene or the expression profile of a large number of related or unrelated 20 genes. When the expression of a single gene is examined, arrays are employed to detect the expression of a specific gene or its variants. When an expression profile is examined, arrays provide a platform for identifying genes that are tissue specific, are affected by a substance being tested in a toxicology assay, are part of a signaling cascade, carry out housekeeping functions, or are specifically related to a particular genetic predisposition, 25 condition, disease, or disorder.

Microarrays may be prepared, used, and analyzed using methods known in the art (see for example, Brennan, T.M. et al. (1995) U.S. Patent No. 5,474,796; Schena, M. et al. (1996) Proc. Natl. Acad. Sci. USA 93:10614-10619; Baldeschweiler et al. (1995) PCT 30 application WO95/251116; Shalon, D. et al. (1995) PCT application WO95/35505; Heller, R.A. et al. (1997) Proc. Natl. Acad. Sci. USA 94:21502155; Heller, M.J. et al. (1997) U.S. Patent No. 5,605,662). Various types of microarrays are well known and

thoroughly described in Schena, M., ed. (1999; DNA Microarrays: A Practical Approach, Oxford University Press, London).

In further embodiments, oligonucleotides or longer fragments derived from any of the 5 polynucleotides described herein may be used as elements on a microarray. The microarray can be used in transcript imaging techniques which monitor the relative expression levels of large numbers of genes simultaneously as described below. The microarray may also be used to identify genetic variants, mutations, and polymorphisms. This information may be used to determine gene function, to understand the genetic 10 basis of a disorder, to diagnose a disorder, to monitor progression/regression of disease as a function of gene expression, and to develop and monitor the activities of therapeutic agents in the treatment of disease. In particular, this information may be used to develop a pharmacogenomic profile of a patient in order to select the most appropriate and effective treatment regimen for that patient. For example, therapeutic 15 agents which are highly effective and display the fewest side effects may be selected for a patient based on his/her pharmacogenomic profile.

As determined by Microarray analysis, DG931 shows differential expression in human primary adipocytes. A strong downregulation is observed concerning the expression of 20 DG931 is during the human adipocyte differentiation (see FIGURE 3). The DG931 protein in preadipocytes has the potential to inhibit adipose differentiation at a very early stage. Therefore, the DG931 protein might play an essential role in adipogenesis. The results are suggesting a role of DG931 in the regulation in human metabolism, for example, as effector / modulator (for example, inhibitor) of adipogenesis. Thus, DG931 25 is a strong candidate for the manufacture of a pharmaceutical composition and a medicament for the treatment of conditions related to human metabolism, such as obesity, diabetes, and/or metabolic syndrome.

The invention also encompasses polynucleotides that encode the proteins of the 30 invention and homologous proteins. Accordingly, any nucleic acid sequence, which encodes the amino acid sequences of the proteins of the invention and homologous proteins, can be used to generate recombinant molecules that express the proteins of the

invention and homologous proteins. In a particular embodiment, the invention encompasses a nucleic acid encoding DG931. It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of nucleotide sequences encoding the proteins, some bearing minimal homology to the nucleotide sequences of any known and naturally occurring gene, may be produced. The invention contemplates each and every possible variation of nucleotide sequence that can be made by selecting combinations based on possible codon choices.

Also encompassed by the invention are polynucleotide sequences that are capable of hybridizing to the claimed nucleotide sequences, and in particular, those of the polynucleotide encoding the proteins of the invention, under various conditions of stringency. Hybridization conditions are based on the melting temperature (T_m) of the nucleic acid binding complex or probe, as taught in Wahl & Berger (1987: Methods Enzymol. 152:399-407) and Kimmel (1987; Methods Enzymol. 152:507-511), and may be used at a defined stringency. Preferably, hybridization under stringent conditions means that after washing for 1 h with 1 x SSC and 0.1% SDS at 50°C, preferably at 55°C, more preferably at 62°C and most preferably at 65°C, particularly for 1 h in 0.2 x SSC and 0.1% SDS at 50°C, preferably at 55°C, more preferably at 62°C and most preferably at 65°C, a positive hybridization signal is observed. Altered nucleic acid sequences encoding the proteins which are encompassed by the invention include deletions, insertions or substitutions of different nucleotides resulting in a polynucleotide that encodes the same or a functionally equivalent protein.

The encoded proteins may also contain deletions, insertions or substitutions of amino acid residues, which produce a silent change and result in functionally equivalent proteins. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues as long as the biological activity of the protein is retained. Furthermore, the invention relates to peptide fragments of the proteins or derivatives thereof such as cyclic peptides, retro-inverso peptides or peptide mimetics having a length of at least 4, preferably at least 6 and up to 50 amino acids.

Also included within the scope of the present invention are alleles of the genes encoding the proteins of the invention and homologous proteins. As used herein, an 'allele' or 'allelic sequence' is an alternative form of the gene, which may result from at least one mutation in the nucleic acid sequence. Alleles may result in altered mRNAs or 5 polypeptides whose structures or function may or may not be altered. Any given gene may have none, one or many allelic forms. Common mutational changes, which give rise to alleles, are generally ascribed to natural deletions, additions or substitutions of nucleotides. Each of these types of changes may occur alone or in combination with the others, one or more times in a given sequence.

10

The nucleic acid sequences encoding DG931 and homologous proteins may be extended utilizing a partial nucleotide sequence and employing various methods known in the art to detect upstream sequences such as promoters and regulatory elements.

15 In order to express a biologically active protein, the nucleotide sequences encoding the proteins or functional equivalents, may be inserted into appropriate expression vectors, i.e., a vector which contains the necessary elements for the transcription and translation of the inserted coding sequence. Methods, which are well known to those skilled in the art, may be used to construct expression vectors containing sequences encoding the 20 proteins and the appropriate transcriptional and translational control elements. Regulatory elements include for example a promoter, an initiation codon, a stop codon, a mRNA stability regulatory element, and a polyadenylation signal. Expression of a polynucleotide can be assured by (i) constitutive promoters such as the Cytomegalovirus (CMV) promoter/enhancer region, (ii) tissue specific promoters such as the insulin 25 promoter (see, Soria et al., 2000, Diabetes 49:157), SOX2 gene promotor (see Li et al., (1998) Curr. Biol. 8:971-974), Msi-1 promotor (see Sakakibara et al., (1997) J. Neuroscience 17:8300-8312), alpha-cardia myosin heavy chain promotor or human atrial natriuretic factor promotor (Klug et al., (1996) J. clin. Invest 98:216-224; Wu et al., (1989) J. Biol. Chem. 264:6472-6479) or (iii) inducible promoters such as the 30 tetracycline inducible system. Expression vectors can also contain a selection agent or marker gene that confers antibiotic resistance such as the neomycin, hygromycin or puromycin resistance genes. These methods include in vitro recombinant DNA

techniques, synthetic techniques, and in vivo genetic recombination. Such techniques are described in Sambrook, J. et al. (1989) *Molecular Cloning. A Laboratory Manual*, Cold Spring Harbor Press, Plainview, N.Y. and Ausubel, F.M. et al. (1989) *Current Protocols in Molecular Biology*, John Wiley & Sons, New York, N.Y.

5

In a further embodiment of the invention, natural, modified or recombinant nucleic acid sequences encoding the proteins of the invention and homologous proteins may be ligated to a heterologous sequence to encode a fusion protein.

- 10 A variety of expression vector/host systems, as known in the art, may be utilized to contain and express sequences encoding the proteins or fusion proteins. These include, but are not limited to, micro-organisms such as bacteria transformed with recombinant bacteriophage, plasmid or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with virus expression vectors (e.g.,
- 15 baculovirus, adenovirus, adeno-associated virus, lentivirus, retrovirus); plant cell systems transformed with virus expression vectors (e.g., cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or PBR322 plasmids); or animal cell systems.
- 20 The presence of polynucleotide sequences of the invention in a sample can be detected by DNA-DNA or DNA-RNA hybridization and/or amplification using probes or portions or fragments of said polynucleotides. Nucleic acid amplification based assays involve the use of oligonucleotides or oligomers based on the sequences specific for the gene to detect transformants containing DNA or RNA encoding the corresponding protein. As
- 25 used herein 'oligonucleotides' or 'oligomers' refer to a nucleic acid sequence of at least about 10 nucleotides and as many as about 60 nucleotides, preferably about 15 to 30 nucleotides, and more preferably about 20-25 nucleotides, which can be used as a probe or amplimer.
- 30 A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting polynucleotide sequences include

oligo-labeling, nick translation, end-labeling of labeled RNA probes, PCR amplification using a labeled nucleotide, or enzymatic synthesis. These procedures may be conducted using a variety of commercially available kits (Pharmacia & Upjohn, (Kalamazoo, Mich.); Promega (Madison Wis.); and U.S. Biochemical Corp., (Cleveland, Ohio).

5

The presence of DG931 in a sample can be determined by immunological methods or activity measurement. A variety of protocols for detecting and measuring the expression of proteins, using either polyclonal or monoclonal antibodies specific for the protein or reagents for determining protein activity are known in the art. Examples include 10 enzyme-linked immunosorbent assay (ELISA), radioimmunoassay (RIA), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on the protein is preferred, but a competitive binding assay may be employed. These and other assays are described, among other places, in Hampton, R. et al. (1990; Serological Methods, a 15 Laboratory Manual, APS Press, St Paul, Minn.) and Maddox, D. E. et al. (1983; J. Exp. Med. 158:1211-1216).

20 Suitable reporter molecules or labels, which may be used, include radionuclides, enzymes, fluorescent, chemiluminescent or chromogenic agents as well as substrates, co-factors, inhibitors, magnetic particles, and the like.

25 The nucleic acids encoding the proteins of the invention can be used to generate transgenic animal or site specific gene modifications in cell lines. Transgenic animals may be made through homologous recombination, where the normal locus of the genes encoding the proteins of the invention is altered. Alternatively, a nucleic acid construct is randomly integrated into the genome. Vectors for stable integration include plasmids, retrovirusses and other animal viruses, YACs, and the like. The modified cells or animal are useful in the study of the function and regulation of the proteins of the invention. For example, a series of small deletions and/or substitutions may be made in 30 the genes that encode the proteins of the invention to determine the role of particular domains of the protein, functions in pancreatic differentiation, etc.

Specific constructs of interest include anti-sense molecules, which will block the expression of the proteins of the invention, or expression of dominant negative mutations. A detectable marker, such as for example lac-Z, may be introduced in the locus of the genes of the invention, where upregulation of expression of the genes of the invention will result in an easily detected change in phenotype.

One may also provide for expression of the genes of the invention or variants thereof in cells or tissues where it is not normally expressed or at abnormal times of development. In addition, by providing expression of the proteins of the invention in cells in which they are not normally produced, one can induce changes in cell behavior.

DNA constructs for homologous recombination will comprise at least portions of the genes of the invention with the desired genetic modification, and will include regions of homology to the target locus. DNA constructs for random integration need not include regions of homology to mediate recombination. Conveniently, markers for positive and/or negative selection are included. Methods for generating cells having targeted gene modifications through homologous recombination are known in the art. For non-human embryonic stem (ES) cells, an ES cell line may be employed, or embryonic cells may be obtained freshly from a host, e.g. mouse, rat, guinea pig etc. Such cells are grown on an appropriate fibroblast-feeder layer or grown in presence of leukemia inhibiting factor (LIF).

When non-human ES or embryonic cells or somatic pluripotent stem cells have been transformed, they may be used to produce transgenic animals. After transformation, the cells are plated onto a feeder layer in an appropriate medium. Cells containing the construct may be detected by employing a selective medium. After sufficient time for colonies to grow, they are picked and analyzed for the occurrence of homologous recombination or integration of the construct. Those colonies that are positive may then be used for embryo transfection and blastocyst injection. Blastocysts are obtained from 4 to 6 week old superovulated females. The ES cells are trypsinized, and the modified cells are injected into the blastocoel of the blastocyst. After injection, the blastocysts are returned to each uterine horn of pseudopregnant females. Females are then allowed to go

to term and the resulting offspring screened for the construct. By providing for a different phenotype of the blastocyst and the genetically modified cells, chimeric progeny can be readily detected. The chimeric animals are screened for the presence of the modified gene and males and females having the modification are mated to produce homozygous progeny. If the gene alterations cause lethality at some point in development, tissues or organs can be maintained as allogenic or congenic grafts or transplants, or in vitro culture. The transgenic animals may be any non-human mammal, such as laboratory animal, domestic animals, etc. The transgenic animals may be used in functional studies, drug screening, etc.

10

The data disclosed in this invention show that the DG931 nucleic acids and proteins and effector / modulator molecules thereof are useful in diagnostic and therapeutic applications implicated, for example, but not limited to, metabolic syndrome, obesity, diabetes mellitus, eating disorder, cachexia, hypertension, coronary heart disease, hypercholesterolemia (dyslipidemia), and/or gallstones. Hence, diagnostic and therapeutic uses for the proteins of the invention nucleic acids and proteins of the invention are, for example but not limited to, the following: (i) ~~protein therapy~~, (ii) small molecule drug target, (iii) antibody target (therapeutic, diagnostic, drug targeting/cytotoxic antibody), (iv) diagnostic and/or prognostic marker, (v) gene therapy (gene delivery/gene ablation), (vi) research tools, and (vii) tissue regeneration in vitro and in vivo (regeneration for all these tissues and cell types composing these tissues and cell types derived from these tissues).

25 The DG931 nucleic acids and proteins and effectors / modulators thereof are useful in diagnostic and therapeutic applications implicated in various applications as described below. For example, but not limited to, cDNAs encoding the proteins of the invention and particularly their human homologues may be useful in gene therapy, and the proteins of the invention and particularly their human homologues may be useful when administered to a subject in need thereof. By way of non-limiting example, the compositions of the present invention will have efficacy for treatment of patients suffering from, for example, obesity, diabetes, and/or metabolic syndrome as described above.

The nucleic acids of the invention or fragments thereof, may further be useful in diagnostic applications, wherein the presence or amount of the nucleic acids or the proteins are to be assessed. Further antibodies that bind immunospecifically to the novel 5 substances of the invention may be used in therapeutic or diagnostic methods.

For example, in one aspect, antibodies, which are specific for the proteins of the invention and homologous proteins, may be used directly as an effector / modulator, e.g. an antagonist or indirectly as a targeting or delivery mechanism for bringing a 10 pharmaceutical agent to cells or tissue which express the protein. The antibodies may be generated using methods that are well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric single chain, Fab fragments, and fragments produced by a Fab expression library. Neutralising antibodies, (i.e., those which inhibit dimer formation) are especially preferred for therapeutic use.

15 For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans, and others, may be immunized by injection with the protein or any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. It is preferred that 20 the peptides, fragments or oligopeptides used to induce antibodies to the protein have an amino acid sequence consisting of at least five amino acids, and more preferably at least 10 amino acids.

Monoclonal antibodies to the proteins may be prepared using any technique that 25 provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique (Köhler, G. et al. (1975) Nature 256:495-497; Kozbor, D. et al. (1985) J. Immunol. Methods 81:31-42; Cote, R. J. et al. Proc. Natl. Acad. Sci. 80:2026-2030; Cole, S. P. et al. (1984) Mol. Cell Biol. 30 62:109-120).

In addition, techniques developed for the production of 'chimeric antibodies', the splicing

of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity can be used (Morrison, S. L. et al. (1984) Proc. Natl. Acad. Sci. 81:6851-6855; Neuberger, M. S. et al (1984) Nature 312:604-608; Takeda, S. et al. (1985) Nature 314:452-454). Alternatively, techniques described for the

5 production of single chain antibodies may be adapted, using methods known in the art, to produce single chain antibodies specific for the proteins of the invention and homologous proteins. Antibodies with related specificity, but of distinct idiotypic composition, may be generated by chain shuffling from random combinatorial immunoglobulin libraries (Burton, D. R. (1991) Proc. Natl. Acad. Sci. 88:11120-11123).

10 Antibodies may also be produced by inducing *in vivo* production in the lymphocyte population or by screening recombinant immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature (Orlandi, R. et al. (1989) Proc. Natl. Acad. Sci. 86:3833-3837; Winter, G. et al. (1991) Nature 349:293-299).

15 Antibody fragments, which contain specific binding sites for the proteins may also be generated. For example, such fragments include, but are not limited to, the $F(ab')_2$ fragments which can be produced by Pepsin digestion of the antibody molecule and the Fab fragments which can be generated by reducing the disulfide bridges of $F(ab')_2$ fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and

20 easy identification of monoclonal Fab fragments with the desired specificity (Huse, W. D. et al. (1989) Science 254:1275-1281).

25 Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding and immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between the protein and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering protein epitopes are preferred, but a competitive binding assay may also

30 be employed (Maddox, *supra*).

In another embodiment of the invention, the polynucleotides or fragments thereof or

nucleic acid ~~effector~~ / modulator molecules such as antisense molecules, aptamers, RNAi molecules or ribozymes may be used for therapeutic purposes. In one aspect, aptamers, i.e. nucleic acid molecules, which are capable of binding to a protein of the invention and modulating its activity, may be generated by a screening and selection 5 procedure involving the use of combinatorial nucleic acid libraries.

- In a further aspect, antisense molecules may be used in situations in which it would be desirable to block the transcription of the mRNA. In particular, cells may be transformed with sequences complementary to polynucleotides encoding DG931 and homologous 10 proteins. Thus, antisense molecules may be used to modulate ~~effector~~ protein activity or to achieve regulation of gene function. Such technology is now well known in the art, and sense or antisense oligomers or larger fragments, can be designed from various locations along the coding or control regions of sequences encoding the proteins. Expression vectors derived from retroviruses, adenovirus, herpes or vaccinia viruses or 15 from various bacterial plasmids may be used for delivery of nucleotide sequences to the targeted organ, tissue or cell population. Methods, which are well known to those skilled in the art, can be used to construct recombinant vectors, which will express antisense molecules complementary to the polynucleotides of the genes encoding the proteins of the invention and homologous proteins. These techniques are described both in 20 Sambrook et al. (supra) and in Ausubel et al. (supra). Genes encoding the proteins of the invention and homologous proteins can be turned off by transforming a cell or tissue with expression vectors, which express high levels of polynucleotides that encode the proteins of the invention and homologous proteins or fragments thereof. Such constructs may be used to introduce untranslatable sense or antisense sequences into a cell. Even in 25 the absence of integration into the DNA, such vectors may continue to transcribe RNA molecules until they are disabled by endogenous nucleases. Transient expression may last for a month or more with a non-replicating vector and even longer if appropriate replication elements are part of the vector system.
- 30 As mentioned above, modifications of gene expression can be obtained by designing antisense molecules, e.g. DNA, RNA or nucleic acid analogues such as PNA, to the control regions of the genes encoding DG931 and homologous proteins, i.e., the

promoters, enhancers, and introns. Oligonucleotides derived from the transcription initiation site, e.g., between positions -10 and +10 from the start site, are preferred. Similarly, inhibition can be achieved using "triple helix" base-pairing methodology.

Triple helix pairing is useful because it cause inhibition of the ability of the double helix

5 to open sufficiently for the binding of polymerases, transcription factors or regulatory molecules. Recent therapeutic advances using triplex DNA have been described in the literature (Gee, J. E. et al. (1994) In: Huber, B. E. and B. I. Carr, Molecular and Immunologic Approaches, Futura Publishing Co., Mt. Kisco, N.Y.). The antisense molecules may also be designed to block translation of mRNA by preventing the

10 transcript from binding to ribosomes.

Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by

15 endonucleolytic cleavage. Examples, which may be used, include engineered hammerhead motif ribozyme molecules that can be specifically and efficiently catalyze endonucleolytic cleavage of sequences encoding the proteins of the invention and homologous proteins. Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites which

20 include the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides corresponding to the region of the target gene containing the cleavage site may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with

25 complementary oligonucleotides using ribonuclease protection assays.

Nucleic acid ~~effector/modulator~~ molecules, e.g. antisense molecules and ribozymes may be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for chemically synthesizing oligonucleotides such

30 as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by in vitro and in vivo transcription of DNA sequences. Such DNA sequences may be incorporated into a variety of vectors with suitable RNA polymerase

promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize antisense RNA constitutively or inducibly can be introduced into cell lines, cells or tissues. RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences 5 at the 5' and/or 3' ends of the molecule or modifications in the nucleobase, sugar and/or phosphate moieties, e.g. the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of non-traditional bases such as inosine, queosine, and wybutosine, as well as 10 acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

Many methods for introducing vectors into cells or tissues are available and equally suitable for use in vivo, in vitro, and ex vivo. For ex vivo therapy, vectors may be 15 introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection and by liposome injections may be achieved using methods, which are well known in the art. Any of the therapeutic methods described above may be applied to any suitable subject including, for example, mammals such as dogs, cats, cows, horses, rabbits, monkeys, and most 20 preferably, humans.

An additional embodiment of the invention relates to the administration of a pharmaceutical composition, in conjunction with a pharmaceutically acceptable carrier, for any of the therapeutic effects discussed above. Such pharmaceutical compositions 25 may consist of DG931 nucleic acids and the proteins and homologous nucleic acids or proteins, antibodies to the proteins of the invention and homologous proteins, mimetics, agonists, antagonists or inhibitors of the proteins of the invention and homologous proteins or nucleic acids. The compositions may be administered alone or in combination with at least one other agent, such as stabilizing compound, which may be 30 administered in any sterile, biocompatible pharmaceutical carrier, including, but not limited to, saline, buffered saline, dextrose, and water. The compositions may be administered to a patient alone or in combination with other agents, drugs or hormones.

The pharmaceutical compositions utilized in this invention may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual or rectal means.

5

In addition to the active ingredients, these pharmaceutical compositions may contain suitable pharmaceutically-acceptable carriers comprising excipients and auxiliaries, which facilitate processing of the active compounds into preparations, which can be used pharmaceutically. Further details on techniques for formulation and administration
10 may be found in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing Co., Easton, Pa.).

Pharmaceutical compositions suitable for use in the invention include compositions
15 wherein the active ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art. For any compounds, the therapeutically effective ~~dosage~~ can be estimated initially either in cell culture assays, e.g., of preadipocyte cell lines or in animal models, usually mice, rabbits, dogs or pigs. The animal model may also be used
20 to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans. A therapeutically effective dose refers to that amount of active ingredient, for example the DG931 nucleic acids or proteins or fragments thereof or antibodies, which is sufficient for treating a specific condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or experimental
25 animals, e.g., ED50 (the dose therapeutically effective in 50% of the population) and LD50 (the dose lethal to 50% of the population). The dose ratio between therapeutic and toxic effects is the therapeutic index, and it can be expressed as the ratio, LD50/ED50. Pharmaceutical compositions, which exhibit large therapeutic indices, are preferred. The data obtained from cell culture assays and animal studies is used in formulating a range
30 of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that include the ED50 with little or no toxicity. The dosage varies within this range depending upon the dosage form employed,

sensitivity of the patient, and the route of administration. The exact dosage will be determined by the practitioner, in light of factors related to the subject that requires treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors, which may be taken into 5 account, include the severity of the disease state, general health of the subject, age, weight, and gender of the subject, diet, time and frequency of administration, drug combination(s), reaction sensitivities, and tolerance/response to therapy. Long-acting pharmaceutical compositions may be administered every 3 to 4 days, every week or once every two weeks depending on half-life and clearance rate of the particular formulation.

10 Normal dosage amounts may vary from 0.1 to 100,000 microg, up to a total dose of about 1 g, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or 15 polypeptides will be specific to particular cells, conditions, locations, etc.

In another embodiment, antibodies which specifically bind to the proteins may be used for the diagnosis of conditions or diseases characterized by or associated with over- or underexpression of the proteins of the invention and homologous proteins or in assays to 20 monitor patients being treated with the proteins of the invention and homologous proteins, or ~~effectors~~ // ~~modulators~~ thereof, e.g. agonists, antagonists, or inhibitors. Diagnostic assays include methods which utilize the antibody and a label to detect the protein in human body fluids or extracts of cells or tissues. The antibodies may be used with or without modification, and may be labeled by joining them, either covalently or 25 non-covalently, with a reporter molecule. A wide variety of reporter molecules, which are known in the art may be used several of which are described above.

A variety of protocols including ELISA, RIA, and FACS for measuring proteins are known in the art and provide a basis for diagnosing altered or abnormal levels of gene 30 expression. Normal or standard values for gene expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, preferably human, with antibodies to the protein under conditions suitable for complex formation. The

amount of standard complex formation may be quantified by various methods, but preferably by photometric means. Quantities of protein expressed in control and disease, samples e.g. from biopsied tissues are compared with the standard values. Deviation between standard and subject values establishes the parameters for diagnosing disease.

5

In another embodiment of the invention, the polynucleotides specific for the DG931 proteins and homologous proteins may be used for diagnostic purposes. The polynucleotides, which may be used, include oligonucleotide sequences, antisense RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect and 10 quantitate gene expression in biopsied tissues in which gene expression may be correlated with disease. The diagnostic assay may be used to distinguish between absence, presence, and excess gene expression, and to monitor regulation of protein levels during therapeutic intervention.

15 In one aspect, hybridization with probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding the proteins of the invention and homologous proteins or closely related molecules, may be used to identify nucleic acid sequences which encode the respective protein. The hybridization probes of the subject invention may be DNA or RNA and are preferably derived from the nucleotide sequence 20 of the polynucleotide encoding the proteins of the invention or from a genomic sequence including promoter, enhancer elements, and introns of the naturally occurring gene. Hybridization probes may be labeled by a variety of reporter groups, for example, radionuclides such as ^{32}P or ^{35}S or enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

25

Polynucleotide sequences specific for DG931 proteins and homologous nucleic acids may be used for the diagnosis of conditions or diseases, which are associated with the expression of the proteins. Examples of such diseases include the obesity, diabetes, metabolic syndrome, and/or others). Polynucleotide sequences specific for the DG931 30 proteins and homologous proteins may also be used to monitor the progress of patients receiving treatment for obesity, diabetes, and/or metabolic syndrome. The polynucleotide sequences may be used qualitative or quantitative assays, e.g. in Southern

or Northern analysis, dot blot or other membrane-based technologies; in PCR technologies; or in dip stick, pin, ELISA or chip assays utilizing fluids or tissues from patient biopsies to detect altered gene expression.

5 In a particular aspect, the DG931 nucleotide sequences may be useful in assays that detect activation or induction of various metabolic diseases or dysfunctions. The nucleotide sequences may be labeled by standard methods, and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is

10 quantitated and compared with a standard value. The presence of altered levels of nucleotide sequences encoding the proteins of the invention and homologous proteins in the sample indicates the presence of the associated disease. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials or in monitoring the treatment of an individual patient.

15 In order to provide a basis for the diagnosis of a disease associated with expression of the DG931 proteins and homologous proteins, a normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence or a

20 fragment thereof, which is specific for the nucleic acids encoding the proteins of the invention and homologous nucleic acids, under conditions suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with those from an experiment where a known amount of a substantially purified polynucleotide is used. Standard values obtained from normal

25 samples may be compared with values obtained from samples from patients who are symptomatic for disease. Deviation between standard and subject values is used to establish the presence of disease. Once disease is established and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to evaluate whether the level of expression in the patient begins to approximate that, which is observed in the

30 normal patient. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

With respect to obesity, diabetes, and/or metabolic syndrome, the presence of an unusual amount of transcript in biopsied tissue from an individual may indicate a predisposition for the development of the disease or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this 5 type may allow health professionals to employ preventative measures or aggressive treatment earlier thereby preventing the development or further progression of the metabolic diseases and disorders.

Additional diagnostic uses for oligonucleotides designed from the sequences encoding 10 the proteins of the invention and homologous proteins may involve the use of PCR. Such oligomers may be chemically synthesized, generated enzymatically or produced from a recombinant source. Oligomers will preferably consist of two nucleotide sequences, one with sense orientation (5prime.fwdarw.3prime) and another with antisense (3prime.rarw.5prime), employed under optimized conditions for identification of a 15 specific gene or condition. The same two oligomers, nested sets of oligomers or even a degenerate pool of oligomers may be employed under less stringent conditions for detection and/or quantification of closely related DNA or RNA sequences.

In another embodiment of the invention, the nucleic acid sequences may also be used to 20 generate hybridization probes, which are useful for mapping the naturally occurring genomic sequence. The sequences may be mapped to a particular chromosome or to a specific region of the chromosome using well known techniques. Such techniques include FISH, FACS or artificial chromosome constructions, such as yeast artificial chromosomes, bacterial artificial chromosomes, bacterial P1 constructions or single 25 chromosome cDNA libraries as reviewed in Price, C. M. (1993) Blood Rev. 7:127-134, and Trask, B. J. (1991) Trends Genet. 7:149-154. FISH (as described in Verma et al. (1988) Human Chromosomes: A Manual of Basic Techniques, Pergamon Press, New York, N.Y.). The results may be correlated with other physical chromosome mapping techniques and genetic map data. Examples of genetic map data can be found in the 30 1994 Genome Issue of Science (265:1981f). Correlation between the location of the gene encoding the proteins of the invention on a physical chromosomal map and a specific disease or predisposition to a specific disease, may help to delimit the region of

DNA associated with that genetic disease.

The nucleotide sequences of the subject invention may be used to detect differences in gene sequences between normal, carrier or affected individuals. An analysis of 5 polymorphisms, e.g. single nucleotide polymorphisms may be carried out. Further, in situ hybridization of chromosomal preparations and physical mapping techniques such as linkage analysis using established chromosomal markers may be used for extending 10 genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may reveal associated markers even if the number or arm of a particular human chromosome is not known. New sequences can be assigned to 15 chromosomal arms or parts thereof, by physical mapping. This provides valuable information to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once the disease or syndrome has been crudely localized by genetic linkage to a particular genomic region, for example, AT to 11q22-23 (Gatti, R. A. et al. (1988) *Nature* 336:577-580), any sequences mapping to that area may represent 20 associated or regulatory genes for further investigation. The nucleotide sequences of the subject invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc. among normal, carrier or affected individuals.

25 In another embodiment of the invention, the proteins of the invention, their catalytic or immunogenic fragments or oligopeptides thereof, an in vitro model, a genetically altered cell or animal, can be used for screening libraries of compounds in any of a variety of drug screening techniques. One can identify effectors, e.g. receptors, enzymes, proteins, ligands, or substrates that bind to, modulate or mimic the action of one or more of the DG931 proteins of the invention. The protein or fragment thereof employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes, between the DG931 proteins 30 of the invention and the agent tested, may be measured. Agents could also, either directly or indirectly, influence the activity of the proteins of the invention.

In addition activity of the proteins of the invention against their physiological substrate(s) or derivatives thereof could be measured in cell-based assays. Agents may

also interfere with posttranslational modifications of the protein, such as phosphorylation and dephosphorylation, farnesylation, palmitoylation, acetylation, alkylation, ubiquitination, proteolytic processing, subcellular localization and degradation. Moreover, agents could influence the dimerization or oligomerization of the proteins of the invention or, in a heterologous manner, of the proteins of the invention with other proteins, for example, but not exclusively, docking proteins, enzymes, receptors, or translation factors. Agents could also act on the physical interaction of the proteins of this invention with other proteins, which are required for protein function, for example, but not exclusively, their downstream signaling.

10 Methods for determining protein-protein interaction are well known in the art. For example binding of a fluorescently labeled peptide derived from the interacting protein to the DG931 protein of the invention, or vice versa, could be detected by a change in polarisation. In case that both binding partners, which can be either the full length proteins as well as one binding partner as the full length protein and the other just represented as a peptide are fluorescently labeled, binding could be detected by fluorescence energy transfer (FRET) from one fluorophore to the other. In addition, a variety of commercially available assay principles suitable for detection of protein-protein interaction are well known in the art, for example but not exclusively

15 AlphaScreen (PerkinElmer) or Scintillation Proximity Assays (SPA) by Amersham. Alternatively, the interaction of the DG931 proteins of the invention with cellular proteins could be the basis for a cell-based screening assay, in which both proteins are fluorescently labeled and interaction of both proteins is detected by analysing cotranslocation of both proteins with a cellular imaging reader, as has been developed

20 for example, but not exclusively, by Cellomics or EvotecOAI. In all cases the two or more binding partners can be different proteins with one being the protein of the invention, or in case of dimerization and/or oligomerization the protein of the invention itself.

25 30 Of particular interest are screening assays for agents that have a low toxicity for mammalian cells. The term "agent" as used herein describes any molecule, e.g. protein or pharmaceutical, with the capability of altering or mimicking the physiological

function of one or more of the proteins of the invention. Candidate agents encompass numerous chemical classes, though typically they are organic molecules, preferably small organic compounds having a molecular weight of more than 50 and less than about 2,500 Daltons. Candidate agents comprise functional groups necessary for

5 structural interaction with proteins, particularly hydrogen bonding, and typically include at least an amine, carbonyl, hydroxyl or carboxyl group, preferably at least two of the functional chemical groups. The candidate agents often comprise carbocyclic or heterocyclic structures and/or aromatic or polyaromatic structures substituted with one or more of the above functional groups.

10

Candidate agents are also found among biomolecules including peptides, saccharides, fatty acids, steroids, purines, pyrimidines, nucleic acids and derivatives, structural analogs or combinations thereof. Candidate agents are obtained from a wide variety of sources including libraries of synthetic or natural compounds. For example, numerous 15 means are available for random and directed synthesis of a wide variety of organic compounds and biomolecules, including expression of randomized oligonucleotides and oligopeptides. Alternatively, libraries of natural compounds in the form of bacterial, fungal, plant and animal extracts are available or readily produced. Additionally, natural or synthetically produced libraries and compounds are readily modified through 20 conventional chemical, physical and biochemical means, and may be used to produce combinatorial libraries. Known pharmacological agents may be subjected to directed or random chemical modifications, such as acylation, alkylation, esterification, amidification, etc. to produce structural analogs. Where the screening assay is a binding assay, one or more of the molecules may be joined to a label, where the label can 25 directly or indirectly provide a detectable signal.

Another technique for drug screening, which may be used, provides for high throughput screening of compounds having suitable binding affinity to the protein of interest as described in published PCT application WO84/03564. In this method, as applied to the 30 proteins of the invention large numbers of different small test compounds, e.g. aptamers, peptides, low-molecular weight compounds etc., are provided or synthesized on a solid substrate, such as plastic pins or some other surface. The test compounds are reacted

with the proteins or fragments thereof, and washed. Bound proteins are then detected by methods well known in the art. Purified proteins can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support. In 5 another embodiment, one may use competitive drug screening assays in which neutralizing antibodies capable of binding the protein specifically compete with a test compound for binding the protein. In this manner, the antibodies can be used to detect the presence of any peptide, which shares one or more antigenic determinants with the protein.

10

The nucleic acids encoding the DG931 proteins of the invention can be used to generate transgenic cell lines and animals. These transgenic non-human animals are useful in the study of the function and regulation of the proteins of the invention *in vivo*. Transgenic animals, particularly mammalian transgenic animals, can serve as a 15 model system for the investigation of many developmental and cellular processes common to humans. A variety of non-human models of metabolic disorders can be used to test modulators of the protein of the invention. Misexpression (for example, overexpression or lack of expression) of the protein of the invention, particular feeding conditions, and/or administration of biologically active compounds can create 20 models of metabolic disorders.

In one embodiment of the invention, such assays use mouse models of insulin resistance and/or diabetes, such as mice carrying gene knockouts in the leptin pathway (for example, ob (leptin) or db (leptin receptor) mice), as described above. 25 In addition to testing the expression of the proteins of the invention in such mouse strains (see EXAMPLES), these mice could be used to test whether administration of a candidate modulator alters for example lipid accumulation in the liver, in plasma, or adipose tissues using standard assays well known in the art, such as FPLC, colorimetric assays, blood glucose level tests, insulin tolerance tests and others.

30

Transgenic animals may be made through homologous recombination in embryonic stem cells, where the normal locus of the gene encoding the protein of the invention

is mutated. Alternatively, a nucleic acid construct encoding the protein is injected into oocytes and is randomly integrated into the genome. One may also express the genes of the invention or variants thereof in tissues where they are not normally expressed or at abnormal times of development. Furthermore, variants of the genes 5 of the invention like specific constructs expressing anti-sense molecules or expression of dominant negative mutations, which will block or alter the expression of the proteins of the invention may be randomly integrated into the genome. A detectable marker, such as lac Z or luciferase may be introduced into the locus of the genes of the invention, where upregulation of expression of the genes of the 10 invention will result in an easily detectable change in phenotype. Vectors for stable integration include plasmids, retroviruses and other animal viruses, yeast artificial chromosomes (YACs), and the like. DNA constructs for homologous recombination will contain at least portions of the genes of the invention with the desired genetic modification, and will include regions of homology to the target locus. Conveniently, 15 markers for positive and negative selection are included. DNA constructs for random integration do not need to contain regions of homology to mediate recombination. DNA constructs for random integration will consist of the nucleic acids encoding the proteins of the invention, a regulatory element (promoter), an intron and a polyadenylation signal. Methods for generating cells having targeted gene modifications 20 through homologous recombination are known in the field. For embryonic stem (ES) cells, an ES cell line may be employed, or embryonic cells may be obtained freshly from a host, e.g. mouse, rat, guinea pig, etc. Such cells are grown on an appropriate fibroblast-feeder layer and are grown in the presence of leukemia inhibiting factor (LIF). ES or embryonic cells may be transfected and can then be used to produce 25 transgenic animals. After transfection, the ES cells are plated onto a feeder layer in an appropriate medium. Cells containing the construct may be selected by employing a selection medium. After sufficient time for colonies to grow, they are picked and analyzed for the occurrence of homologous recombination. Colonies that are positive may then be used for embryo manipulation and morula aggregation. Briefly, morulae 30 are obtained from 4 to 6 week old superovulated females, the Zona Pellucida is removed and the morulae are put into small depressions of a tissue culture dish. The ES cells are trypsinized, and the modified cells are placed into the depression closely

to the morulae. On the following day the aggregates are transferred into the uterine horns of pseudopregnant females. Females are then allowed to go to term. Chimeric offsprings can be readily detected by a change in coat color and are subsequently screened for the transmission of the mutation into the next generation (F1-generation). Offspring of the F1-generation are screened for the presence of the modified gene and males and females having the modification are mated to produce homozygous progeny. If the gene alterations cause lethality at some point in development, tissues or organs can be maintained as allogenic or congenic grafts or transplants, or in vitro culture. The transgenic animals may be any non-human mammal, such as laboratory animal, domestic animals, etc., for example, mouse, rat, guinea pig, sheep, cow, pig, and others. The transgenic animals may be used in functional studies, drug screening, and other applications and are useful in the study of the function and regulation of the proteins of the invention in vivo.

15 Finally, the invention also relates to a kit comprising at least one of

- (a) a nucleic acid molecule coding for a protein of the invention or a functional fragment thereof;
- (b) a protein of the invention or a fragment or an isoform thereof;
- (c) a vector comprising the nucleic acid of (a);

20 (d) a host cell comprising the nucleic acid of (a) or the vector of (b);

- (e) a polypeptide encoded by the nucleic acid of (a);
- (f) a fusion polypeptide encoded by the nucleic acid of (a);
- (g) an antibody, an aptamer or another effector / modulator against the nucleic acid of (a) or the polypeptide of (b), (e) or (f) and

25 (h) an anti-sense oligonucleotide of the nucleic acid of (a).

The kit may be used for diagnostic or therapeutic purposes or for screening applications as described above. The kit may further contain user instructions.

Figure 1 shows human DG931 protein

Figure 1A shows the nucleic acid sequence of human DG931 protein (SEQ ID NO: 1)

Figure 1B shows the amino acid sequence (one-letter code) of human DG931 protein (SEQ ID NO: 2)

5

Figure 2 shows the analysis of DG931 protein expression in mammalian tissues. The relative RNA-expression is shown on the Y-axis, in Figure 2A to 2C the tissues tested are given on the X-axis. WAT refers to white adipose tissue, BAT refers to brown adipose tissue. In Figure 2D, the X-axis represents the time axis. 'd0' refers to day 0 (start of the experiment), 'd2' - 'd10' refers to day 2 - day 10 of adipocyte differentiation).

10 **Figure 2A shows the quantitative analysis of DG931 expression in mouse wild-type tissues.**

15 **Figure 2B shows the quantitative analysis of DG931 expression in wild-type mice (wt-mice), compared to genetically obese mice (ob/ob-mice) and to fasted mice (fasted-mice).**

20 **Figure 2C shows the quantitative analysis of DG931 expression in mice fed with a control diet compared to mice fed with a high fat diet.**

Figure 2D shows the quantitative analysis of DG931 expression in mammalian fibroblast (3T3-L1) cells, during the differentiation from preadipocytes to mature adipocytes.

Figure 3 shows the quantitative analysis of DG931 expression in human abdominal adipocyte cells, during the differentiation from preadipocytes to mature adipocytes.

25

The examples illustrate the invention:

Example 1: Identification of secreted factors expressed in pancreas

30 **A screen for secreted factors expressed in developing mouse pancreas was carried out according to methods known by those skilled in the art (see, for example Pera E.M. and De Robertis E.M., (2000) Mech Dev 96(2): 183-195) with several**

modifications.

Expression cDNA library:

5 A mouse embryonic stage 9.5-15 pancreatic bud library was prepared in pCMVSPORT-6 vector using SUPERSCRIPT Plasmid System from Invitrogen according to the manufacturer's instructions. The non-amplified library was electroporated into MaxEff DH10B cells (Invitrogen).

Secretion cloning

10 Bacterial clones were picked with sterile toothpicks from agar plates and cultured in 96-deep-well microtiter plates in LB-ampicillin (see Sambrook et al., supra). Aliquots of 8 cultures were pooled, and plasmid DNA was isolated using the BioRobot_9600 apparatus according to the manufactures instructions (Qiagen; QIAprep® Turbo BioRobot Kit. Human 293 cell culture cells were cultured in 75 ml
15 tissue culture flasks in DMEM and 10% fetal calf serum. At 90-99% confluence, the cells were splitted at 1:3 ratio and plated onto poly-D-lysine (Sigma) coated 96-well plates. Cells were transfected with 100-500 ng plasmid using lipofectamine 2000 (Invitrogen). After 6 hours, the medium was exchanged for fresh complete growth medium. 24 hours after transfection, the cells were washed twice with DMEM
20 without cysteine and methionine (Invitrogen), supplemented with 1% dialysed Bovine serum (Sigma) with 50 microgram per ml Heparin (Sigma) and glutamine. The cells were labeled radioactively ' (S35 Met-label', from Hartmann Analytic GmbH). After 12 hours, aliquots of the supernatants were harvested in 96-well PCR plates and subjected to SDS gel electrophoresis in precast 4±20% gradient
25 polyacrylamide Criterion gels (Biorad) under reducing conditions, using Criterion Dodeca Cell gel running chamber (Biorad). The gels were fixed in 10% acetic acid, 25% isopropanol for 30 min, soaked 15-30 min in AMPLIFY reagent (Amersham), dried and exposed to X-OMAT (AR) film (Kodak). Positive clones were identified and regrown in 96-well-plates. DNA of individual clones was prepared and used for
30 transfection as described above. If one of the clones yielded proteins of the same size as that of the original pool, a positive clone was identified. Positive clones were partially sequenced from the 5prime end (SEQLAB, Goettingen).

Example 2: Identification of the human DG931 homologous proteins (FIGURE 1)

The term "polynucleotide comprising the nucleotide sequence as shown in GenBank

5 Accession number" relates to the expressible gene of the nucleotide sequences deposited under the corresponding GenBank Accession number. The term "GenBank Accession number" relates to NCBI GenBank database entries (Ref.: Benson et al., Nucleic Acids Res. 28 (2000) 15-18).

10 DG931 homologous proteins and nucleic acid molecules coding therefore are obtainable from insect or vertebrate species, e.g. mammals or birds. Particularly preferred are nucleic acids comprising human DG931 homologs. The following mouse sequence was identified in the 'secreted factor screen': *Mus musculus* RIKEN cDNA 1810049K24 gene (1810049K24Rik), GenBank Accession Number NM_030209

15 (2929 base pairs mRNA) and GenBank Accession Number NP_084485 (494 amino acid protein)).

Sequences homologous to mouse DG931 were identified using the publicly available program BLASTP 2.2.3 of the non-redundant protein data base of the National Center for Biotechnology Information (NCBI) (see, Altschul et al., 1997, Nucleic Acids Res. 25:3389-3402). The best human homolog of mouse DG931 is *Homo sapiens* hypothetical protein DKFZp434B044 (DKFZP434B044), GenBank Accession Number NM_031476 (4574 base pairs mRNA; SEQ ID NO: 1) and GenBank Accession Number NP_113664 (497 amino acid protein; SEQ ID NO: 2).

25

Example 3: Expression of the polypeptides in mammalian tissues (FIGURE 2)

To analyse the expression of the polypeptides disclosed in this invention in mammalian tissues, several mouse strains (preferably mice strains C57Bl/6J, C57Bl/6 ob/ob and

30 C57Bl/KS db/db which are standard model systems in obesity and diabetes research) were purchased from Harlan Winkelmann (33178 Borch, Germany) and maintained under constant temperature (preferably 22°C), 40 per cent humidity and a light / dark

cycle of preferably 14 / 10 hours. The mice were fed a standard chow (for example, from ssniff Spezialitäten GmbH, order number ssniff M-Z V1126-000). For the fasting experiment ("fasted wild-type mice"), wild-type mice were starved for 48 h without food, but only water supplied ad libitum (see, for example, Schnetzler et al. J Clin Invest 5 1993 Jul;92(1):272-80, Mizuno et al. Proc Natl Acad Sci U S A 1996 Apr 16;93(8):3434-8). Animals were sacrificed at an age of 6 to 8 weeks. The animal tissues were isolated according to standard procedures known to those skilled in the art, snap frozen in liquid nitrogen and stored at -80°C until needed.

10 For analyzing the role of the proteins disclosed in this invention in the in vitro differentiation of different mammalian cell culture cells for the conversion of pre-adipocytes to adipocytes, mammalian fibroblast (3T3-L1) cells (e.g., Green & Kehinde, Cell 1: 113-116, 1974) were obtained from the American Tissue Culture Collection (ATCC, Hanassas, VA, USA; ATCC- CL 173). 3T3-L1 cells were 15 maintained as fibroblasts and differentiated into adipocytes as described in the prior art (e.g., Qiu. et al., J. Biol. Chem. 276:11988-95, 2001; Slieker et al., BBRC 251: 225-9, 1998). In brief, cells were plated in DMEM/10% FCS (Invitrogen, Karlsruhe, Germany) at 50,000 cells/well in duplicates in 6-well plastic dishes and cultured in a humidified atmosphere of 5% CO₂ at 37°C. At confluence (defined as day 0: d0) cells were 20 transferred to serum-free (SF) medium, containing DMEM/HamF12 (3:1; Invitrogen), fetuin (300 µg/ml; Sigma, Munich, Germany), transferrin (2 µg/ml; Sigma), pantothenate (17 µM; Sigma), biotin (1 µM; Sigma), and EGF (0.8nM; Hoffmann-La Roche, Basel, Switzerland). Differentiation was induced by adding dexamethasone (DEX; 1 µM; Sigma), 3-methyl-isobutyl-1-methylxanthine (MIX; 0.5mM; Sigma), and 25 bovine insulin (5 µg/ml; Invitrogen). Four days after confluence (d4), cells were kept in SF medium, containing bovine insulin (5 µg/ml) until differentiation was completed. At various time points of the differentiation procedure, beginning with day 0 (day of confluence) and day 2 (hormone addition; for example, dexamethasone and 3-isobutyl-1-methylxanthine), up to 10 days of differentiation, suitable aliquots of cells 30 were taken every two days.

RNA was isolated from mouse tissues or cell culture cells using Trizol Reagent (for

example, from Invitrogen, Karlsruhe, Germany) and further purified with the RNeasy Kit (for example, from Qiagen, Germany) in combination with an DNase-treatment according to the instructions of the manufacturers and as known to those skilled in the art. Total RNA was reverse transcribed (preferably using Superscript II RNaseH-

5 Reverse Transcriptase, from Invitrogen, Karlsruhe, Germany) and subjected to Taqman analysis preferably using the Taqman 2xPCR Master Mix (from Applied Biosystems, Weiterstadt, Germany; the Mix contains according to the Manufacturer for example AmpliTaq Gold DNA Polymerase, AmpErase UNG, dNTPs with dUTP, passive reference Rox and optimized buffer components) on a GeneAmp 5700 Sequence

10 Detection System (from Applied Biosystems, Weiterstadt, Germany).

The following prime/probe pairs were used for the TaqMan analysis (GenBank Accession Number NM_030209 for the mouse DG931 sequence): Mouse DG931 reverse primer (Seq ID NO:3) 5'-CAGTCATCAACTTCATGACCCAAG -3'; Mouse 15 DG931 reverse primer (Seq ID NO:4) 5'-CTGTTACACGTGGATCCTTG-3'; Mouse DG931 Taqman probe (Seq ID NO:5) (5/6-FAM)-TCCACTGTGACACCAAGATGAAGGACTCA-(5/6-TAMRA)

Example 4. Analysis of the differential expression of transcripts of the proteins of
20 the invention in human tissues

RNA preparation from human primary adipose tissues was done as described in EXAMPLE 2. The hybridization and scanning was performed as described in the manufactures manual (see Affymetrix Technical Manual, 2002, obtained from 25 Affmetrix, Santa Clara, USA).

The expression analysis '(using Affymetrix GeneChips) of the DG931 gene using primary human abdominal adipocyte differentiation clearly shows differential expression of human DG931 in adipocytes. Two independent experiments were done. 30 Both experiments show that the DG931 transcripts are the most abundance at day 0 compared to day 12 during differentiation. These data further confirm the mouse 3T3L-1 differentiation data.

Thus, the DG931 protein has to be significantly decreased in order for the preadipocytes to differentiate into mature adipocyte. The DG931 protein in preadipocytes has the potential to inhibit adipose differentiation. Therefore, the DG931
5 protein might play an essential role in the regulation of human metabolism, in particular in the regulation of adipogenesis and thus it might be an essential role in obesity, diabetes, and/or metabolic syndrome.

For the purpose of the present invention, it will be understood by the person having
10 average skill in the art that all combinations of all individual features mentioned throughout the specification are explicitly disclosed herewith.

Claims

1. A pharmaceutical composition comprising a nucleic acid molecule encoding DG931 protein and / or a functional fragment thereof and an effector / modulator of said nucleic acid molecule and / or said protein or protein fragment.
2. The composition of claim 1, wherein the composition contains pharmaceutically acceptable carriers, diluents, and / or additives.
3. The composition of claim 1 or 2, wherein the nucleic acid molecule is a mammalian DG931 nucleic acid, particularly encoding the human DG931 polypeptide and/or a nucleic molecule which is complementary thereto or a fragment thereof or a variant thereof.
4. The composition of any one of claims 1-3, wherein said nucleic acid molecule is selected from the group consisting of
 - (a) a nucleic acid molecule encoding a polypeptide as shown in SEQ ID NO: 2, or an isoform, fragment or variant of the polypeptide as shown in SEQ ID NO: 2;
 - (b) a nucleic acid molecule which comprises or is the nucleic acid molecule as shown in SEQ ID NO: 1;
 - (c) a nucleic acid molecule being degenerate with as a result of the genetic code to the nucleic acid sequences as defined in (a) or (b),
 - (d) a nucleic acid molecule that hybridizes at 50°C in a solution containing 1 x SSC and 0.1% SDS to a nucleic acid molecule as defined in claim 2 or as defined in (a) to (c) and/or a nucleic acid molecule which is complementary thereto;
 - (e) a nucleic acid molecule that encodes a polypeptide which is at least 85%, preferably at least 90%, more preferably at least 95%, more preferably at least 98% and up to 99,6% identical to the human DG931, as defined in claim 2 or to a polypeptide as defined in (a);
 - (f) a nucleic acid molecule that differs from the nucleic acid molecule of (a) to (e) by mutation and wherein said mutation causes an alteration, deletion, duplication or premature stop in the encoded polypeptide.

5. The composition of any one of claims 1-4, wherein the nucleic acid molecule is a DNA molecule, particularly a cDNA or a genomic DNA.
6. The composition of any one of claims 1-5, wherein said nucleic acid encodes a polypeptide contributing to regulating the metabolism, in particular human metabolism.
7. The composition of any one of claims 1-6, wherein said nucleic acid molecule is a recombinant nucleic acid molecule.
8. The composition of any one of claims 1-7, wherein the nucleic acid molecule is a vector, particularly an expression vector.
9. The composition of any one of claims 1-4, wherein the polypeptide is a recombinant polypeptide.
10. The composition of claim 9, wherein said recombinant polypeptide is a fusion polypeptide.
11. The composition of any one of claims 1-8, wherein said nucleic acid molecule is selected from hybridization probes, primers and anti-sense oligonucleotides.
12. The composition of any one of claims 1-11 which is a diagnostic composition.
13. The composition of any one of claims 1-11 which is a therapeutic composition.
14. The composition of any one of claims 1-13 for the manufacture of an agent for detecting and/or verifying, for the treatment, alleviation and/or prevention of metabolic diseases or dysfunctions, including, obesity, diabetes, and/or metabolic syndrome.
15. Use of a DG931 nucleic acid molecule or a polypeptide encoded thereby or a fragment or a variant of said nucleic acid molecule or said polypeptide and / or an effector / modulator of said

File: DG0101A2

nucleic acid or polypeptide for the manufacture of a medicament for the treatment of obesity, diabetes, and/or metabolic syndrome and/or for controlling the function of a gene and/or a gene product which is influenced and/or modified by a DG931 polypeptide.

16. Use of the DG931 nucleic acid molecule or use of a nucleic acid molecule encoding DG931 or a homologue thereof or use of a polypeptide encoded thereby, or use of a fragment or a variant of said nucleic acid molecule or said polypeptide, or use of an effector / modulator of said nucleic acid molecule or said polypeptide for identifying substances in vitro capable of interacting with a DG931 polypeptide.
17. A non-human transgenic animal exhibiting a modified expression of a DG931 polypeptide.
18. The animal of claim 17, wherein the expression of the DG931 polypeptide is increased or reduced.
19. A recombinant host cell exhibiting a modified expression of a DG931 polypeptide, or a recombinant host cell which comprises a nucleic acid molecule as defined in any one of claims 1 to 8 or 11.
20. The cell of claim 19 which is a human cell.
21. A method of identifying a (poly)peptide involved in the regulation of energy homeostasis and/or metabolism in a mammal comprising the steps of
 - (a) contacting a collection of (poly)peptides with a DG931 homologous polypeptide or a fragment thereof under conditions that allow binding of said (poly)peptides;
 - (b) removing (poly)peptides which do not bind and
 - (c) identifying (poly)peptides that bind to said DG931 homologous polypeptide.
22. A method of screening for an agent which effects / modulates the interaction of a DG931 polypeptide with a binding target comprising the steps of
 - (a) incubating a mixture comprising
 - (aa) a DG931 polypeptide or a fragment thereof;
 - (ab) a binding target/agent of said DG931 polypeptide or fragment thereof;

and

(ac) a candidate agent under conditions whereby said polypeptide or fragment thereof specifically binds to said binding target at a reference affinity;

- (b) detecting the binding affinity of said DG931 polypeptide or fragment thereof to said binding target to determine an affinity for the agent; and
- (c) determining a difference between affinity for the agent and reference affinity.

23. A method for screening for an agent, which modulates the activity of a DG931 polypeptide, comprising the steps of

- (a) incubating a mixture comprising
 - (aa) a DG931 polypeptide or a fragment thereof; and
 - (ab) a candidate agentunder conditions whereby said DG931 polypeptide or fragment thereof exhibits a reference activity,
- (b) detecting the activity of said DG931 polypeptide or fragment thereof to determine a activity for the agent; and
- (c) determining a difference between activity for the agent and reference activity.

24. A method of producing a composition comprising the (poly)peptide identified by the method of claim 21 or the agent identified by the method of claim 22 or 23 with a pharmaceutically acceptable carrier and/or diluent.

25. The method of claim 24 wherein said composition is a pharmaceutical composition for preventing, alleviating or treating of metabolic diseases and disorders, including obesity, diabetes, and/or metabolic syndrome.

26. Use of a (poly)peptide as identified by the method of claim 21 or of an agent as identified by the method of claim 22 or 23 for the preparation of a pharmaceutical composition for the treatment, alleviation and/or prevention of metabolic diseases or dysfunctions, including obesity, diabetes, and/or metabolic syndrome.

27. Use of a nucleic acid molecule as defined in any one of claims 1 to 8 or 11 for the preparation of a medicament for the treatment, alleviation and/or prevention of metabolic diseases or dysfunctions, including obesity, diabetes, and/or metabolic syndrome.

28. Use of a polypeptide as defined in any one of claims 1 to 6, 9 or 10 for the preparation of a medicament for the treatment, alleviation and/or prevention of metabolic diseases or dysfunctions, including obesity, diabetes, and/or metabolic syndrome.
29. Use of a vector as defined in claim 8 or the preparation of a medicament for the treatment, alleviation and/or prevention of metabolic diseases or dysfunctions, including obesity, diabetes, and/or metabolic syndrome.
30. Use of a host cell as defined in claim 19 or 20 for the preparation of a medicament for the treatment, alleviation and/or prevention of metabolic diseases or dysfunctions, including obesity, diabetes, and/or metabolic syndrome.
31. Use of a DG931 nucleic acid molecule or of a fragment thereof for the production of a non-human transgenic animal which over- or under-expresses the DG931 gene product.
32. Kit comprising at least one of
 - (a) a DG931 nucleic acid molecule or a functional fragment or an isoform thereof;
 - (b) a DG931 amino acid molecule or a functional fragment or an isoform thereof;
 - (c) a vector comprising the nucleic acid of (a);
 - (d) a host cell comprising the nucleic acid of (a) or the vector of (b);
 - (e) a polypeptide encoded by the nucleic acid of (a), expressed by the vector of (c) or the host cell of (a);
 - (f) a fusion polypeptide encoded by the nucleic acid of (a);
 - (g) an antibody, an aptamer or another effector / modulator against the nucleic acid of (a) or the polypeptide of (b), (e), or (f) and /or
 - (h) an anti-sense oligonucleotide of the nucleic acid of (a).

31. März 2003

Abstract

The present invention discloses proteins secreted by the developing pancreas, and polynucleotides, which identify and encode these proteins. The invention also relates to the use of these sequences in the diagnosis, study, prevention, and treatment of metabolic diseases and disorders.

Figure 1. Nucleic acid and amino acid sequences of human DG931

Figure 1A. Human DG931, nucleic acid sequence (SEQ ID NO:1)

1 ggccggctgtc cccattgagc tgcgtctgtcg ctgtgcccgc tgcgtgcgtc gtgcccgcgc
61 tgcgtccgtc gctaccgcgt ctgcgtggacg cgggagacgc cagcgcgcgtg gtgattggag
121 ccctgcggag agctcaagcg cccagctctg cccgaggagc ccaggtgcgc cctgtgatcc
181 catagttgtc gcaggagtgg agccatgagc tgcgtctgtgg gtgggtgtcat ccccttgggg
241 ctgcgtttcc tggctctgcgg atcccaaggc tacctcctgc ccaacgtcac tctcttagag
301 gagctgctca gcaaataccca gcacaacgag ttcactccc gggtccgcag agccatcccc
361 agggaggaca aggaggagat cctcatgctg cacaacaagc ttccgggcca ggtgcagcct
421 caggcctcca acatggagta catgacactgg gatgacgaac tggagaagtc tgctgcagcg
481 tggggccagtc agtgcacatctg ggagcacggg cccaccagtc tgctgggtgtc catcgggcag
541 aacctggcg ctcactgggg caggtatcgc ttcctggggt tccatgtgca gtcctggtat
601 gacgaggtga aggactacac ctacccctac ccgagcgcgt gcaacccctg gtgtccagag
661 aggtgctcg ggccatgtg cacgcactac acacagatag tttggggcac caccaacaag
721 atcggttgtg ctgtgaacac ctgcggaaag atgactgtct ggggagaagt ttgggagaac
781 gcggtctact ttgtctgcaa ttattctcca aaggggaact ggattggaga agccccctac
841 aagaatggcc ggccctgctc tgagtgcaca cccagctatg gaggcagctg caggaacaac
901 ttgtgttacc gagaagaaac ctacactcca aaacctgaaa cggacgcgt gaatgagggt
961 gaaacggctc ccattcctga agaaaaccat gtttggctcc aaccgagggt gatgagaccc
1021 accaagccca agaaaaccctc tgccgtcaac tacatgaccc aagtgcgtcag atgtgacacc
1081 aagatgaagg acaggtgcaa agggtccacg tgtaacaggt accagtgcac agcaggctgc
1141 ctgaaccaca aggcgaagat ctttggaaact ctgttctatg aaagtcgtc tagcatatgc
1201 cgcgcgcgcca tccactacgg gatcctggat gacaaggagg gcctgggtgg tatcaccagg
1261 aacgggaagg tcccccttctt cgtgaagtct gagagacacg gcgtgcagtc cctcagcaaa
1321 tacaaacccctt ccagctcatt catgggtgtca aaagtgaaaag tgcaggattt ggactgctac
1381 acgaccgttg ctcagctgtg cccgtttgaa aagccagcaa ctcactgcac aagaatccat
1441 tgcgtccggcac actgcaaaaga cgaacccctcc tactgggtctc cgggtttgg acccaacatc
1501 tatgcagata cctcaagcat ctgcgtacaca gcccgtcagc cgggagtcgt cagcaacgag
1561 agtgggggtg acgtggacgt gatgcccgtg gataaaaaga agacctacgt gggctcgctc
1621 aggaatggag ttcaactgtga aagcctgggg actcctcggtt atggaaaggc cttccggatc
1681 ttgtgtgtca ggcagtgaat ttccagcacc aggggagaag gggcgttttc aggagggctt
1741 cggggttttt cttttatccc tattttgtca ttgcggggta tatggagagt caggaaactt

1801 cctttgactg atgttcagtg tccatcactt tgtggcctgt gggtgaggtg acatctcatc
1861 ccctcactga agcaacagca tcccaagggtg ctcagccgga ctccctggtg cctgatcctg
1921 ctggggcccg ggggtctcca tctggacgtc ctctctcctt tagagatctg agctgtctct
1981 taaaggggac agttccccaa aatgttcctt gctatgtgtt ctctgttgg tggaggaagt
2041 tgatttcaac ctccctgcca aaagaacaaa ccatttgaag ctcacaattt tgaagcattc
2101 acggcgctgg aagaggcctt ttgagcaagc gccaatgagt ttcaaggaatg aagtagaaagg
2161 tagttattta aaaataaaaa acacagtccg tccctaccaa tagagaaaaa tggttttat
2221 gtttgcgtt cagacagaca aatgggctag agtaagaggg ctgcgggtat gagagacccc
2281 ggctccgccc tggcacgtgt ccttgcgtgc ggcccggccac aggccccctt caatggccgc
2341 attcaggatg gctctataca cagcagtgtct ggtttatgtt gaggcagca gtcacttcag
2401 agatgtatct tgcgttgc aggccttca tcttcatggc ccacctgttt tctgcgtgt
2461 cctttggtcc cattgaggac taaggatcgg gaccctttctt ttacccctta cccattgtgg
2521 ctccccacctt gcctcggact ggtttacgtt tcttgcgttca caccaggac ttttctttgc
2581 aagcgaacct gtttgaagcc caagtcttaa ctccctggtct cgtaagggttc cactgagacg
2641 agatgtctga gaacaaccaa agaaggcctg ctcttgcgtt ctttaaaaaa atgacaattt
2701 aatgtgcaga ttccccacgc acccgatgac ctatttttc agccgtggga ggaatggagt
2761 ctttggtaca ttccctcaccg aggttagcag ctcaaggatgtt ggttatgaaa ccgtctgtgg
2821 cctcatgaca gcgagagatg ggaatacact agaaggatct ctttccctgt tttcgtgaaa
2881 cgactcttgc caaacgttcc cgaggcgcca aggagtgttag tacaccctgg ctgccatcac
2941 tctataaaaag tgcttcatga gccagacca aaagccccaca gtgaaatgaa gtaccctttt
3001 gtaaatagca ttttttgca gaagggtaaa attccactct ctaccaccgg gccagccaat
3061 agatcacttt ggtgaatgtt agtttcaaat ttgattcaaa atatttctta ggtgaaagaa
3121 cttagcagaaa gtcaaaaact aagatactgt agactggaca agaaattcta cctgggcacc
3181 taggtgatgc cttcttctt tgattgcctt tctaataaat gcagaatctg aaggtaaata
3241 gtttaaaaac aaaacaaaaaa cccaccctt taaggagttg gtaaaaagca gtcacttct
3301 tagcttgact gagctaaaat tcacaggact acgtgccttgc tgcattgttag tctagtcgtt
3361 attcataggt actgactcct cagccccaaa tgcggagag gaagaattcg gtcagcctgt
3421 cagggtcgta gtccaggatc caccacat ctggggaaact tctgggtgtt gggtgctctg
3481 ctgcgtggact tttgtggctg tgcgtgtgtc tgcaagataa attagatgc cctgtgggggt
3541 ttgcagaatt agtgaagggtt ccaggacat cccagttgggc tgcgttccaa agatccac
3601 tcaaggaga cttgaaactt ccagtgtgag ttgacccat cattttaaaaaa taaagtcccc
3661 gggttccctta atgcctcctt cactgggcct tccctagcagg atagaaagtc cttgcccaga
3721 gcaggacctg gtcgtctttttttt tttcccgaga ccaagttca ctctgttgcc
3781 caaggttagag tgcagtggcg tgcgtctgc tcaatttgcac tgcgcctcc cgggttcaag

3841 caattctcat gcatcagcct cccaaagtacc tgggactaca ggcgtgagct accatgccccg
3901 gctaattttt gtatTTTtag tagagatggg gtttattat gttggccagg ctggtctcga
3961 actccttacc tcaggtgatc caccacattt ggccctccga agtgctggga ttacaggcat
4021 gagccactgc gcccggccat ggacctggct gtctttatca tccccacaaa cattttgaaa
4081 ctggaatatt tggatTTcaga aaatggaaac aagactataa atgataagcc ctgtccctag
4141 caccacctct cctgtgtgtg gaatagaggc ccctcgtgct accaacaactt accctgtgtt
4201 taaaaaagata ttgtaccaag ccaacggcgt tcctggctct cctgcccaca ggatgaacat
4261 ttccggcttc cttaggagtt ttgcctacc gtattccaaa gcgtgtgctg gtttctcata
4321 ttgtctgttag gctcactcag cccgcagttt atgtgtgtgc tttttctat gaaaaatgat
4381 gtattttgtct acttcctgtg tacaaaagttt tattgtaaaat gttttttgtg ctttgcata
4441 acagggggcca cgttgttgca attgtttcag tagaactggt ttgatttcta aatgttcct
4501 gtaacatatac ttttatgaac aaatctgaac aatttgtgaa ataaaacatt gaaaacccaaa
4561 aaaaaaaaaa aaaa

//

Figure 1B. Human DG931, amino acid sequence (SEQ ID NO: 2).

MSCVLGGVPLGLFLVCGSQGYLLPNVTLEELLSKYQHNESH
SRVRAIPREDKEEILMLHNKLRGQVQPQASNMEYMTWDDELEKAAA
WASQCIWEHG
PTSLLVSIGQNLGAHWGRYRSPGFHVQSWYDEVKDYTYP
PSECNPWCPERCSGP
MCT
HYTQIVWATTNKIGCAVN
TCRKMTVWGEVWENAVYFV
CNYS
PKGNWIGEAPYKN
GRPC
SEC
PPSYGGSCRNNLCY
REETYTPK
PETDEMNEV
ETAPI
PEENHV
WLQPRVMRPT
KPK
KTS
AVNYMTQVV
RC
CDT
KM
KDR
CK
G
ST
CN
Y
QCP
AG
CL
NH
KAK
I
G
T
L
F
Y
E
SS
S
C
RA
AIHYG
I
L
DD
K
G
L
V
D
I
T
R
N
G
K
V
P
F
F
V
K
S
E
R
H
G
V
Q
S
L
S
K
Y
K
P
S
S
F
M
V
S
K
V
K
V
Q
D
L
C
Y
TT
V
A
Q
L
C
P
F
E
K
P
A
T
H
C
P
R
I
H
C
P
A
H
C
K
D
E
P
S
Y
W
A
P
V
F
G
T
N
I
Y
A
D
T
S
S
I
C
K
T
A
V
H
A
G
V
I
S
E
G
G
D
V
D
V
M
P
V
D
K
K
T
Y
V
G
S
L
R
N
G
V
Q
S
E
S
L
G
T
P
R
D
G
K
A
F
R
I
F
A
V
R
Q

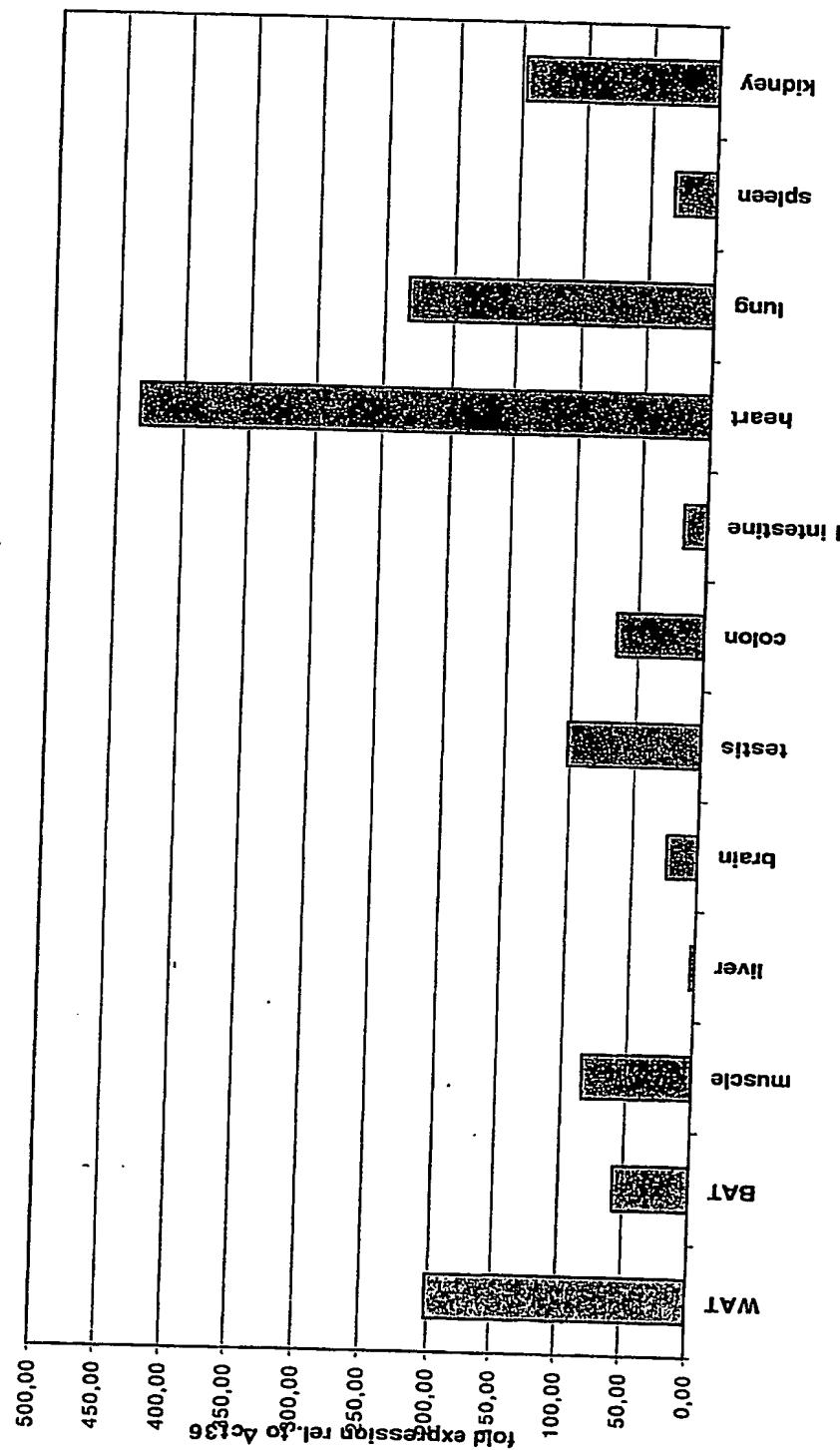


FIGURE 2A. Quantitative analysis of DG931 expression in mouse wild-type tissues

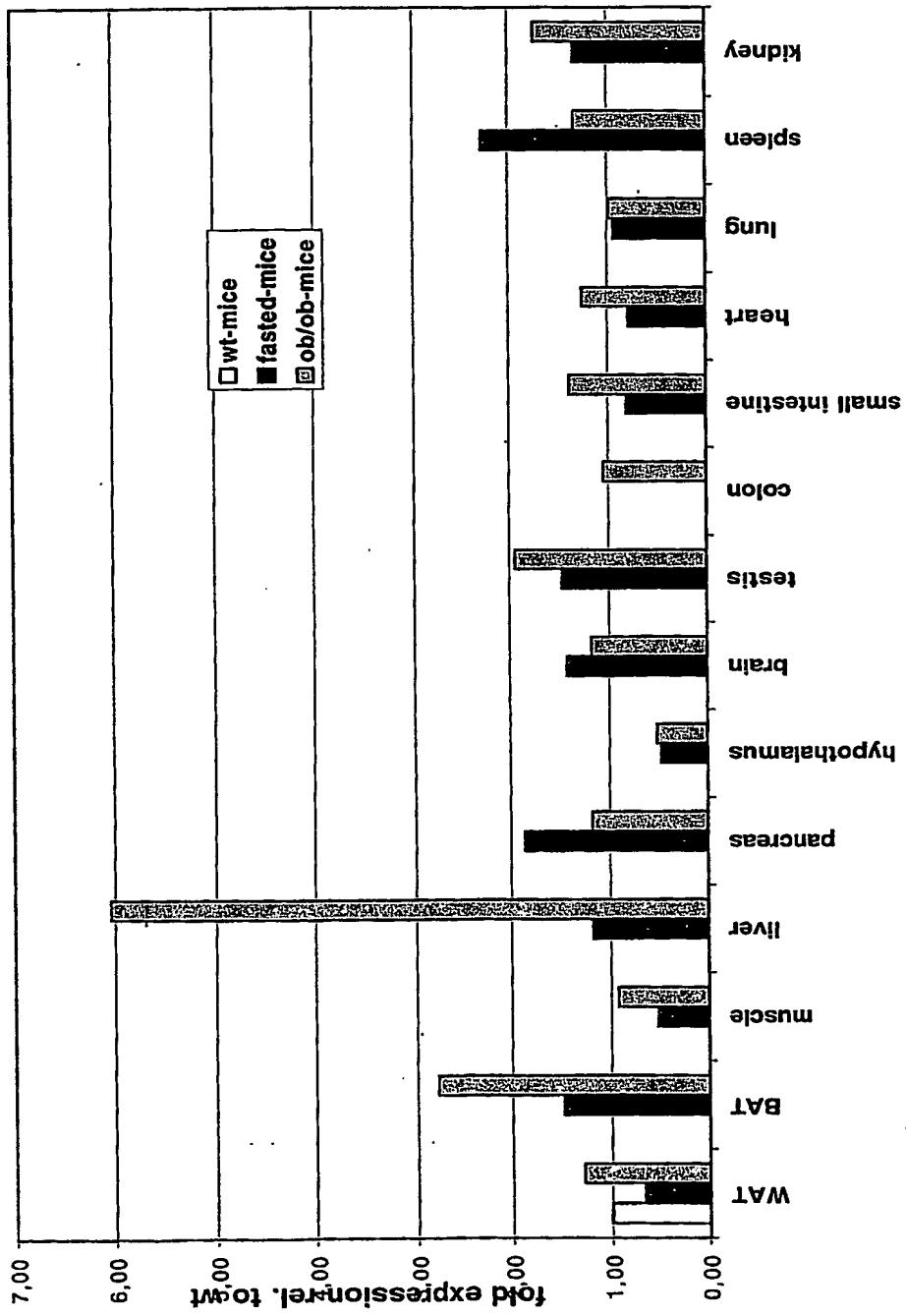


FIGURE 2B. Quantitative analysis of DG931 expression in wild-type mice, compared to genetically obese mice and to fasted mice

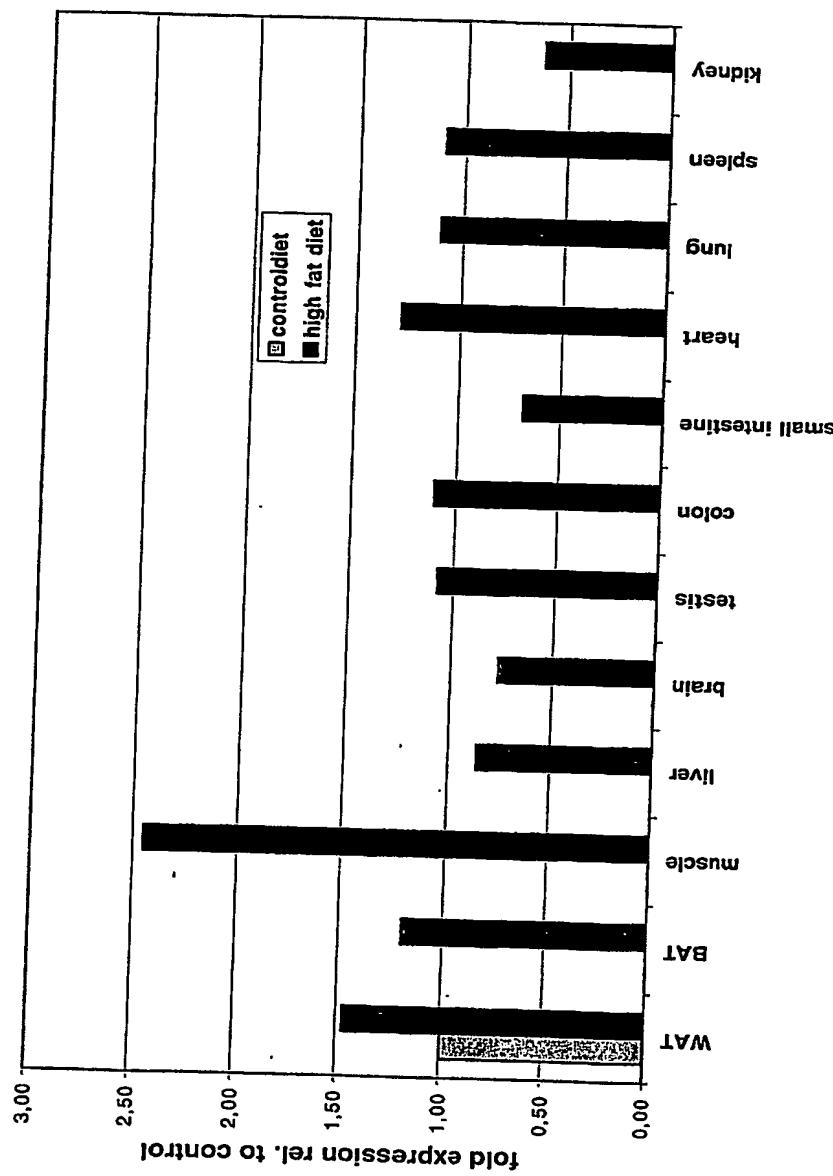
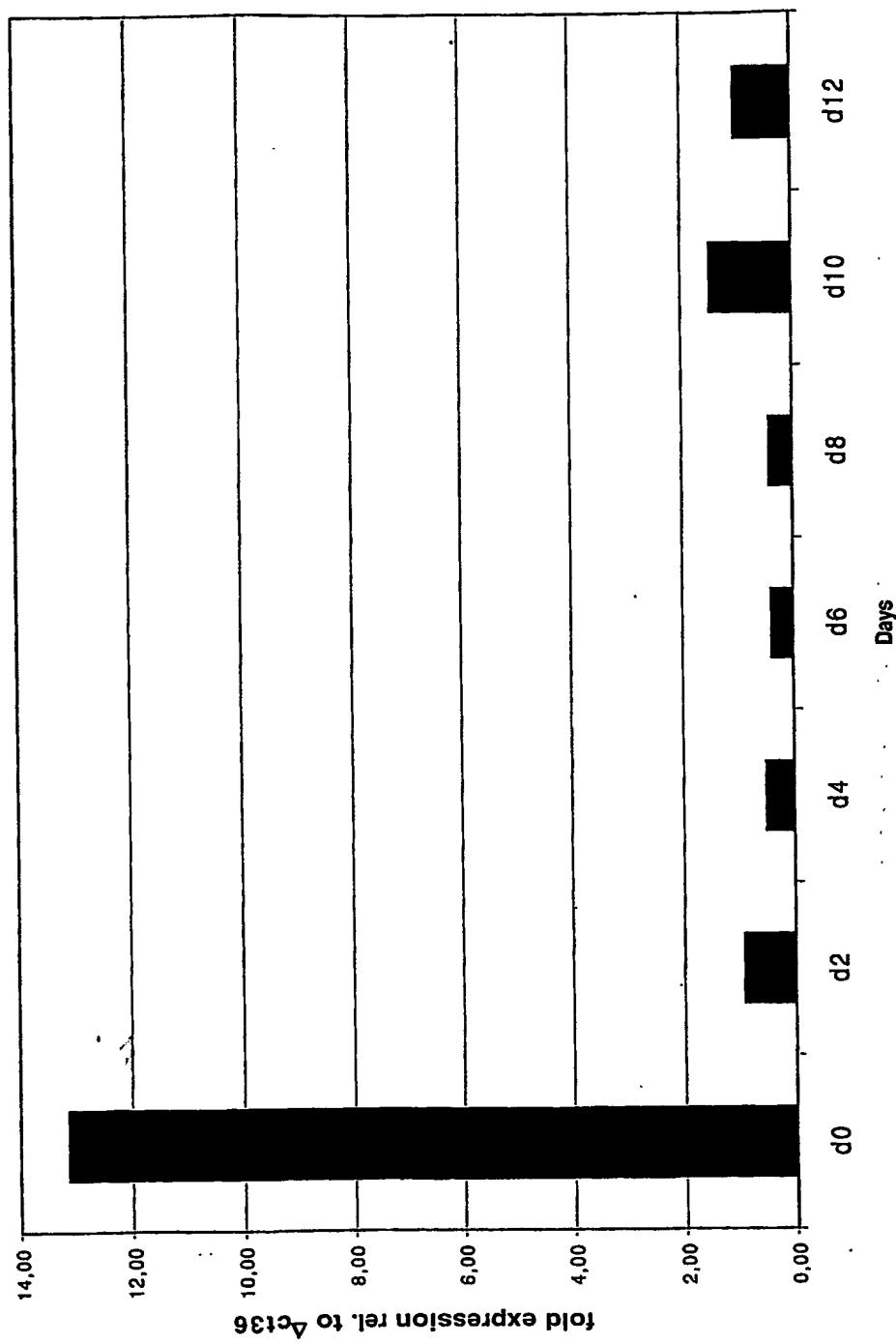


FIGURE 2C. Quantitative analysis of DG931 expression in mice fed with a high fat diet compared to mice fed with a control diet.

FIGURE 2D. Quantitative analysis of DG931 expression in mammalian fibroblast (3T3-L1) cells, during the differentiation from preadipocytes to mature adipocytes.



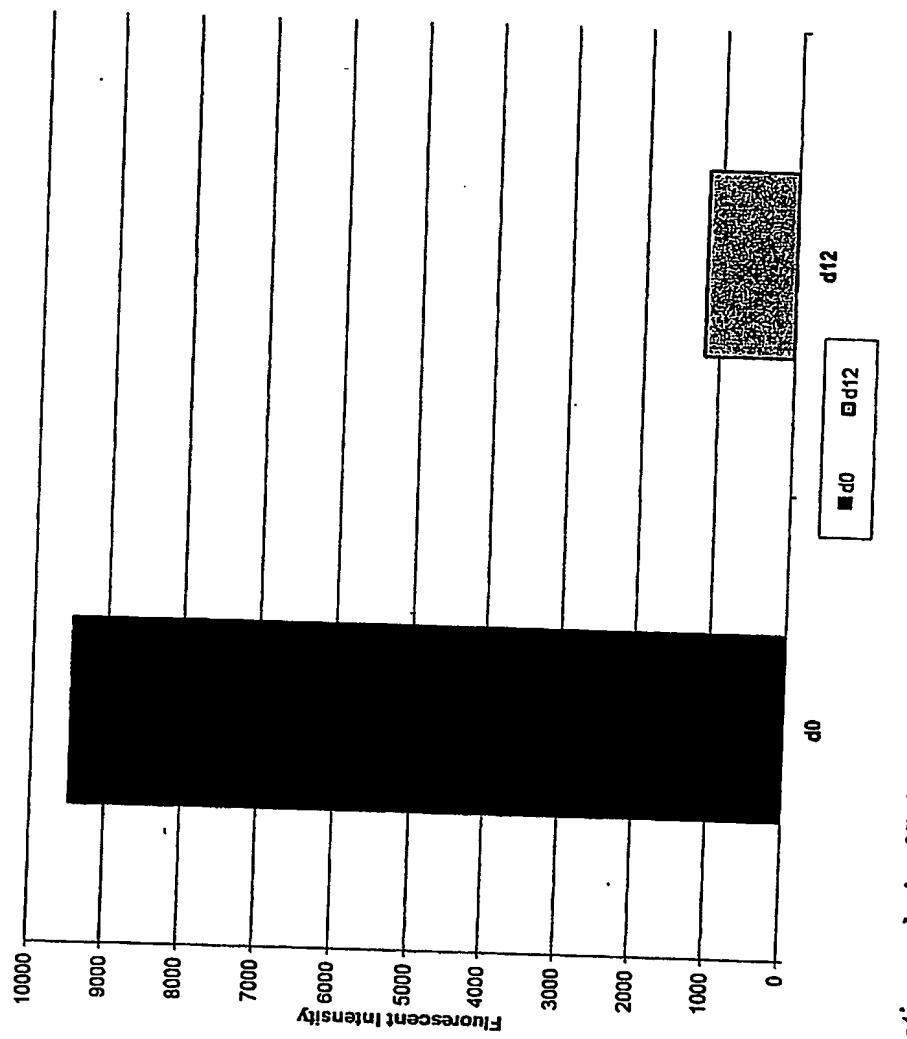


FIGURE 3. quantitative analysis of DG931 expression in human abdominal adipocyte cells, during the differentiation from preadipocytes to mature adipocytes

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- BLACK BORDERS**
- IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- FADED TEXT OR DRAWING**
- BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- SKEWED/SLANTED IMAGES**
- COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- GRAY SCALE DOCUMENTS**
- LINES OR MARKS ON ORIGINAL DOCUMENT**
- REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.